

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-171826
 (43)Date of publication of application : 26.06.1998

(51)Int. Cl. G06F 17/30

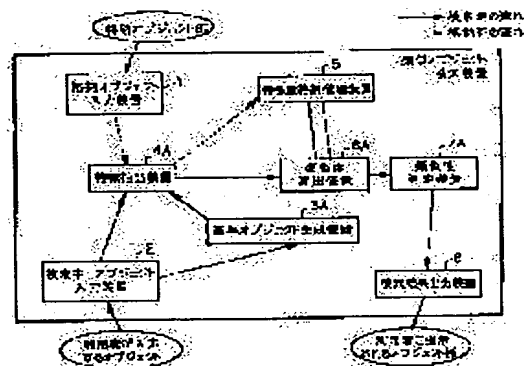
(21)Application number : 08-330813 (71)Applicant : NIPPON TELEGR & TELEPH CORP
 <NTT>
 (22)Date of filing : 11.12.1996 (72)Inventor : AKAMA HIROKI

(54) METHOD FOR RETRIEVING SIMILAR OBJECTS AND DEVICE THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To automatically classify similar objects from dissimilar objects at the time of retrieving similar objects.

SOLUTION: The featured values of a retrieval key object inputted from a retrieval key object inputting device 2 are extracted by a featured value extracting device 4A, and the similarity of both objects is calculated from the featured values of the retrieval key object and the featured values of an input object stored in a featured value storing and managing device 5. A reference object is generated from the retrieval key object by a reference object generating device 3A, the featured values of the reference object are extracted by the featured value extracting device 4A, and similarity α of the both objects is calculated by the similarity calculating device 6A. Afterwards, the objects whose similarity is more than the similarity α are sequenced in the order of the high similarity by a similarity judging device 7A, and outputted in the order of the high similarity by a retrieved result outputting device 8.



LEGAL STATUS

[Date of request for examination]

24.11.2000

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-171826

(43) 公開日 平成10年(1998) 6月26日

(51) Int.Cl.⁶

G 0 6 F 17/30

識別記号

F I

G 0 6 F 15/403

15/40

15/401

3 5 0 C

3 7 0 G

3 1 0 A

審査請求 未請求 請求項の数14 O L (全 14 頁)

(21) 出願番号 特願平8-330813

(22) 出願日 平成8年(1996)12月11日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 赤間 浩樹

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

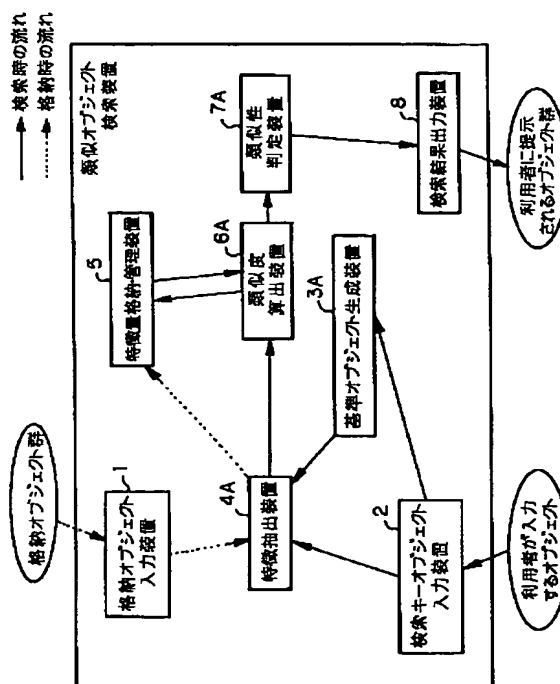
(74) 代理人 弁理士 若林 忠

(54) 【発明の名称】 類似オブジェクト検索方法および装置

(57) 【要約】

【課題】 類似オブジェクトの検索において、類似するオブジェクトと、類似しないオブジェクトを自動的に分類する。

【解決手段】 検索キーオブジェクト入力装置2から入力された検索キーオブジェクトの特徴量が特徴量抽出装置4Aによって抽出され、類似度算出装置6Aによって、検索キーオブジェクトの特徴量と、特徴量格納・管理装置5に格納されている入力オブジェクトの特徴量から両者の類似度を算出する。これと平行して、基準オブジェクト生成装置3Aによって、検索キーオブジェクトから基準オブジェクトが生成され、特徴量抽出装置4Aによってその特徴量が抽出され、類似度算出装置6Aによって両者の類似度 α が算出される。この後、類似性判定装置7Aによって、類似度が類似度 α 以上のオブジェクトについて類似度の高い順に順序付けが行われ、検索結果出力装置8によって類似度の高い順に出力される。



【特許請求の範囲】

【請求項1】 入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、

検索キーオブジェクトを入力する段階と、

前記検索キーオブジェクトの特徴量を抽出する段階と、

前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトと基準オブジェクトの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトとの類似度が前記検索キーオブジェクトと前記基準オブジェクトの類似度以上のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、

順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する類似オブジェクト検索方法。

【請求項2】 入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、

検索キーオブジェクトを入力する段階と、

前記検索キーオブジェクトの特徴量を抽出する段階と、

前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトと基準オブジェクトの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトとの類似度が、前記検索キーオブジェクトと前記基準オブジェクトの類似度以上のオブジェクトを、否定の指定があった場合には前記検索キーオブジェクトとの類似度が、前記検索キーオブジェクトと前記基準オブジェクトの類似度未満のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する類似オブジェクト検索方法。

【請求項3】 入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、

検索キーオブジェクトを入力する段階と、

前記検索キーオブジェクトの特徴量を抽出する段階と、

前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトと複数の基準オブジェクトの類似度を特徴量を使って算出し、それらの類似度の中から所定の選択基準で1つの基準類似度を選択するか、それらの類似度を基にして所定の関数で1つの基準類似度を算出する段階と、

前記検索キーオブジェクトとの類似度が前記基準類似度以上のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、

順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する類似オブジェクト検索方法。

【請求項4】 入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、

検索キーオブジェクトを入力する段階と、

前記検索キーオブジェクトの特徴量を抽出する段階と、

前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、

前記検索キーオブジェクトと複数の基準オブジェクトの類似度を特徴量を使って算出し、それらの類似度の中から所定の選択基準で第1の基準類似度およびこれよりも低い第2の基準類似度を選択するか、前記算出された類似度を基にして所定の関数で第1の基準類似度およびこれよりも低い第2の基準類似度を算出する段階と、

前記検索キーオブジェクトとの類似度が前記第1の基準類似度以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が前記第2の基準類似度以下のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、

順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する類似オブジェクト検索方法。

【請求項5】 前記基準オブジェクトが予め用意されている、請求項1から4のいずれか1項記載の類似オブジェクト検索方法。

【請求項6】 前記基準オブジェクトを、前記検索キーオブジェクトを加工することによって生成する段階を有する、請求項1から4のいずれか1項に記載の類似オブジェクト検索方法。

【請求項7】 前記基準オブジェクトを、前記オブジェクト群を加工または統計処理することによって生成する段階を有する、請求項1から4のいずれか1項記載の類似オブジェクト検索方法。

【請求項8】 格納されるオブジェクトを入力する格納オブジェクト入力手段と、

類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、

前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、

前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、

前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記基準オブジェクトと前記検索キーオブジェクトの類似度を求める類似度算出手段と、

前記検索キーオブジェクトとの類似度が、前記基準オブジェクトと前記検索キーオブジェクトの類似度以上のオ

プロジェクトについてその類似度の大きさにしたがって順序付けを行う類似性判定手段と、
順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する類似オブジェクト検索装置。

【請求項9】 格納されるオブジェクトを入力する格納オブジェクト入力手段と、
類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、
前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、
前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、
前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、
前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記基準オブジェクトと前記検索キーオブジェクトの類似度を求める類似度算出手段と、
前記検索キーオブジェクトとの類似度が、前記基準オブジェクトと前記検索キーオブジェクトの類似度以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が、前記検索キーオブジェクトと前記基準オブジェクトの類似度未満のオブジェクトを、類似度の大きさにしたがって順序付けする類似性判定手段と、
順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する類似オブジェクト検索装置。

【請求項10】 格納されるオブジェクトを入力する格納オブジェクト入力手段と、
類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、
前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、
前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、
前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび複数の基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、
前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度の中から

ら所定の選択基準で1つの基準類似度を選択するか、それらの類似度を基にして所定の関数で1つの基準類似度を算出する類似度算出手段と、
前記検索キーオブジェクトとの類似度が前記基準類似度以上のオブジェクトについて、その類似度の大きさにしたがって順序付けを行う類似性判定手段と、
順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する類似オブジェクト検索装置。

【請求項11】 格納されるオブジェクトを入力する格納オブジェクト入力手段と、
類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、
前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、
前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、
前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび複数の基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、
前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度の中から所定の選択基準で第1の基準類似度およびこれよりも低い第2の基準類似度を選択するか、前記算出された類似度を基にして所定の関数で第1の基準類似度およびこれよりも低い第2の基準類似度を算出する類似度算出手段と、
前記検索キーオブジェクトとの類似度が前記第1の基準類似度以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が前記第2の基準類似度以下のオブジェクトを、類似度の大きさにしたがって順序付けを行う類似性判定手段と、
順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する類似オブジェクト検索装置。

【請求項12】 前記基準オブジェクトが外部から与えられる、請求項8から11のいずれか1項記載の類似オブジェクト検索装置。

【請求項13】 前記検索オブジェクト入力手段によって入力された検索キーオブジェクトを加工することによって前記基準オブジェクトを生成する基準オブジェクト生成手段を有する請求項8から11のいずれか1項記載の類似オブジェクト検索装置。

【請求項14】 前記特徴量格納・管理手段に格納されているオブジェクト群を加工または統計処理することに

前記基準オブジェクトを生成する基準オブジェクト生成手段を有する、請求項8から11のいずれか1項記載の類似オブジェクト検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、利用者が与えたオブジェクトに近いオブジェクトを検索する類似オブジェクト（テキスト、画像、映像、音声、音楽等）の検索方法および装置に関する。

【0002】

【従来の技術】図9は類似オブジェクト検索装置の従来例の構成図、図10はその動作を示すフローチャートである。ここでは主にオブジェクトが画像の場合について述べる。

【0003】類似画像を検索するには、画像特徴量をベクトルとし、類似度としてベクトル空間内の距離を使った以下の方法が一般的である。

【0004】まず、検索対象となる画像の集合を $S = \{S_1, \dots, S_n\}$ 、検索のキーとなる画像を K 、特徴抽出関数を $F(\quad)$ とする。

【0005】画像提供者によって提供される画像 S_i ($1 \leq i \leq n$)は格納オブジェクト入力装置1によって入力され、特徴抽出装置4によって各画像 S_i に対する特徴ベクトル $F(S_i)$ が生成され、この S_i と $F(S_i)$ の組を特徴量格納・管理装置5に格納しておく。この特徴量格納・管理装置5は一般にデータベースを使って容易に実現される。

【0006】類似画像検索装置の利用者が検索キーとなる画像 K を検索キーオブジェクト入力装置2によって入力したとき（ステップ21）、まず、特徴抽出装置4Cによって画像 K の特徴ベクトル $F(K)$ を算出し（ステップ22）、類似度算出装置6Eによって特徴量格納・管理装置5内の全ての特徴ベクトル $F(S_j)$ ($1 \leq j \leq n$)と特徴ベクトル $F(K)$ との距離を求め（ステップ23）、類似性判定装置7Eによって、距離の近い順に類似度が高いと判定し（ステップ24）、検索結果出力装置8によって、その順序付けにしたがって画像 S_j を利用者に提示する（ステップ25）。

【0007】ここで、類似度の基準としてユークリッド距離、内積等の各種の距離基準が用いられる。

【0008】ここで述べた方法は特徴量抽出装置の特徴ベクトルの計算法を交換することによって、画像に限らず、テキスト、音、映像等に適用できる極めて一般的な手法である。

【0009】

【発明が解決しようとする課題】上述した従来の方法では、与えられたキーオブジェクトに類似する順（類似度の高い順）に全てのオブジェクトを並べることが可能であった。しかし、従来の方法では「類似する」と「類似しない」を区切る基準を与えることがなく、全く類似し

ない画像も（低い類似度を持った）類似画像として検索結果に含まれていた。よって利用者は、

・全てのオブジェクトを検索結果とする。

【0010】利用者が検索結果のオブジェクト総数を予め制限する。

【0011】利用者が検索結果を類似度の上位から順に見ながら、適当な時点で打ち切る。

のような対処を行い、「類似する」と「類似しない」の判定は、利用者が自ら行っていた。

【0012】本発明の目的は、「類似するオブジェクト」と「類似しないオブジェクト」とを自動的に分離する類似オブジェクト検索方法および装置を提供することである。

【0013】本発明の他の目的は、否定形の検索、すなわち類似しないオブジェクトの検索を行うことが可能な類似オブジェクト検索方法および装置を提供することにある。

【0014】

【課題を解決するための手段】本発明の第1の類似オブジェクト検索方法は、入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、検索キーオブジェクトを入力する段階と、前記検索キーオブジェクトの特徴量を抽出する段階と、前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトと基準オブジェクトの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトとの類似度が前記検索キーオブジェクトと前記基準オブジェクトの類似度以上のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する。

【0015】本発明の第2の類似オブジェクト検索方法は、入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、検索キーオブジェクトを入力する段階と、前記検索キーオブジェクトの特徴量を抽出する段階と、前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトと基準オブジェクトの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトとの類似度が前記検索キーオブジェクトと前記基準オブジェクトの類似度以上のオブジェクトを、否定の指定があった場合には前記検索キーオブジェクトとの類似度が、前記検索キーオブジェクトと前記基準オブジェクトの類似度未満のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する。

【0016】本発明の第3の類似オブジェクト検索方法は、入力された複数のオブジェクトからそれぞれの特徴

量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、検索キーオブジェクトを入力する段階と、前記検索キーオブジェクトの特徴量を抽出する段階と、前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトと複数の基準オブジェクトの類似度を特徴量を使って算出し、それらの類似度の中から所定の選択基準で1つの基準類似度を選択するか、それらの類似度を基にして所定の関数で1つの基準類似度を算出する段階と、前記検索キーオブジェクトとの類似度が前記基準類似度以上のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する。

【0017】本発明の第4の類似オブジェクト検索方法は、入力された複数のオブジェクトからそれぞれの特徴量を抽出し、オブジェクトと特徴量の対応を記憶手段に格納する段階と、検索キーオブジェクトを入力する段階と、前記検索キーオブジェクトの特徴量を抽出する段階と、前記複数のオブジェクトの前記検索キーオブジェクトとの類似度を特徴量を使って算出する段階と、前記検索キーオブジェクトと複数の基準オブジェクトの類似度を特徴量を使って算出し、それらの類似度の中から所定の選択基準で第1の基準類似度およびこれよりも低い第2の基準類似度を選択するか、前記算出された類似度を基にして所定の関数で第1の基準類似度およびこれよりも低い第2の基準類似度を算出する段階と、前記検索キーオブジェクトとの類似度が前記第1の基準類似度以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が前記第2の基準類似度以下のオブジェクトを、類似度の大きさにしたがって順序付けする段階と、順序付けが行われたオブジェクトを順序にしたがって出力する段階を有する。

【0018】本発明では従来の欠点を解決するため、類似オブジェクトに対する検索の過程において、まず、基準オブジェクトを用意する。その基準オブジェクトと利用者から入力されたオブジェクトとの類似度 α を算出し、その類似度 α を基準値とし、それより類似度が高いオブジェクトを類似するオブジェクトと判定し、それより類似度が低いオブジェクトを類似でないオブジェクトと判定する。

【0019】基準オブジェクトが複数（例えば n 個）存在する場合には、それら全ての基準オブジェクトと検索キーオブジェクトの類似度 α_i ($1 \leq i \leq n$) を算出し、それらの類似度の中でも最も高い類似度 $\beta_1 = \max(\alpha_i)$ 、類似度の中で最も低い類似度 $\beta_2 = \min(\alpha_i)$ を第1の基準類似度、第2の基準類似度とし、第1の基準類似度 β_1 より類似度の高いオブジェクトは類似するオブジェクト、第2の基準類似度 β_2 より類似度の低いオブジェクトは類似でないオブジェクトと判定する。

【0020】なお、第1、第2の基準類似度を求める関数はここでは \max 、 \min としたが、 n 番目に大きい値を求める関数や、平均を求める関数、分散を使った関数等、一般の関数でも構わない。

【0021】また、基準オブジェクトは、予め固定的に用意しておいても構わないし、利用者から入力された検索キーオブジェクトを加工して作成しても構わないし、あるいは、既に格納されているオブジェクト群を加工または統計処理することで生成しても構わない。

【0022】本発明の第1の類似オブジェクト検索装置は、格納されるオブジェクトを入力する格納オブジェクト入力手段と、類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記基準オブジェクトと前記検索キーオブジェクトの類似度を求める類似度算出手段と、前記検索キーオブジェクトとの類似度が、前記基準オブジェクトと前記検索キーオブジェクトの類似度以上のオブジェクトについてその類似度の大きさにしたがって順序付けを行う類似性判定手段と、順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する。

【0023】本発明の第2の類似オブジェクト検索装置は、格納されるオブジェクトを入力する格納オブジェクト入力手段と、類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記基準オブジェクトと前記検索キーオブジェクトの類似度を求める類似度算出手段と、前記検索キーオブジェクトとの類似度が、前記基準オブジェクトと前記検索キーオブジェクトの類似度以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が、前記検索キーオブジェクトと前記基準オブジェクト

の類似度未満のオブジェクトを、類似度の大きさにしたがって順序付けする類似性判定手段と、順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する。

【0024】本発明の第3の類似オブジェクト検索装置は、格納されるオブジェクトを入力する格納オブジェクト入力手段と、類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび複数の基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度の中から所定の選択基準で1つの基準類似度を選択するか、それらの類似度を基にして所定の関数で1つの基準類似度を算出する類似度算出手段と、前記検索キーオブジェクトとの類似度が前記基準類似度以上のオブジェクトについて、その類似度の大きさにしたがって順序付けを行う類似性判定手段と、順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する。

【0025】本発明の第4の類似オブジェクト検索装置は、格納されるオブジェクトを入力する格納オブジェクト入力手段と、類似オブジェクトの検索時、検索キーとなる検索キーオブジェクトを入力する検索キーオブジェクト入力手段と、前記各オブジェクトの特徴量が格納され、これを管理する特徴量格納・管理手段と、前記格納オブジェクト入力手段によって入力された複数のオブジェクトの特徴量を抽出し、該特徴量をオブジェクトと対応づけて前記特徴量格納・管理手段に格納し、前記検索キーオブジェクト入力手段によって入力された検索キーオブジェクトおよび複数の基準オブジェクトからそれぞれの特徴量を抽出する特徴抽出手段と、前記特徴量格納・管理手段に格納されている特徴量を使って前記各オブジェクトの前記検索キーオブジェクトとの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度を求め、前記複数の基準オブジェクトと前記検索キーオブジェクトの類似度の中から所定の選択基準で第1の基準類似度およびこれよりも低い第2の基準類似度を選択するか、前記算出された類似度を基にして所定の関数で第1の基準類似度およびこれよりも低い第2の基準類似度を算出する類似度算出手段と、前記検索キーオブジェクトとの類似度が前記第1の基準類似度

以上のオブジェクトを、否定の指定があった場合は前記検索キーオブジェクトとの類似度が前記第2の基準類似度以下のオブジェクトを、類似度の大きさにしたがって順序付けを行う類似性判定手段と、順序付けが行われたオブジェクトを順序にしたがって出力する検索結果出力手段を有する。

【0026】第1、第2、第3の類似オブジェクト検索装置はそれぞれ第1、第2、第3、第4の類似オブジェクト検索方法に対応している。

【0027】

【発明の実施の形態】次に本発明の実施形態について図面を参照して説明する。

【0028】図1は本発明の第1の実施形態の類似オブジェクト検索装置の構成図、図2はその動作を示すフローチャートである。

【0029】本実施形態の類似オブジェクト検索装置は格納オブジェクト入力装置1と検索キーオブジェクト入力装置2と基準オブジェクト生成装置3Aと特徴抽出装置4Aと特徴量格納・管理装置5と類似度算出装置6Aと類似性判定装置7Aと検索結果出力装置8で構成されている。

【0030】格納オブジェクト入力装置1は、格納されるオブジェクトを入力する。検索キーオブジェクト入力装置2は、類似オブジェクト検索時、検索キーとなる検索キーオブジェクトを入力する。基準オブジェクト生成装置3Aは、検索キーオブジェクト入力装置2によって入力された検索キーオブジェクトを加工して基準オブジェクトを生成する。特徴抽出装置4Aは、格納オブジェクト入力装置1によって入力された複数のオブジェクトの特徴量を抽出し、オブジェクトと対応づけて特徴量格納・管理装置5に格納し、検索キーオブジェクト入力装置2によって入力された検索キーオブジェクトおよび基準オブジェクト生成装置3Aによって生成された基準オブジェクトからそれぞれの特徴量を抽出する。類似度算出装置6Aは特徴量格納・管理装置5に格納されている特徴量を使って各オブジェクトの検索キーオブジェクトとの類似度を求め、基準オブジェクトと検索キーオブジェクトの類似度 α を求める。類似性判定装置7Aは、検索キーオブジェクトとの類似度が、基準オブジェクトと検索キーオブジェクトの類似度 α 以上のオブジェクトについてその類似度の高い順に順序付けを行う。検索結果出力装置8は、順序付けが行われたオブジェクトを類似度の高い順に出力する。

【0031】次に、本実施形態の動作を図2のフローチャートにより説明する。なお、既に複数のオブジェクトが入力され、その特徴量がオブジェクトに対応付けて特徴量格納・管理装置5に格納されているものとする。

【0032】まず、検索キーオブジェクトを検索キーオブジェクト入力装置2から入力する（ステップ11）。すると、特徴量抽出装置4Aによって、検索キーオブジ

ェクトの特徴量が抽出され（ステップ12）、類似度算出装置6Aによって、検索キーオブジェクトの特徴量と、特徴量格納・管理装置5に格納されている入力オブジェクトの特徴量から検索キーオブジェクトと各入力オブジェクトの類似度を算出する（ステップ13）。これと平行して、基準オブジェクト生成装置3Aによって、検索キーオブジェクトから基準オブジェクトが生成され、特徴量抽出装置4Aによってその特徴量が抽出され（ステップ14A）、類似度算出装置6Aによって、検索キーオブジェクトの特徴量と基準オブジェクトの特徴量から両者の類似度 α が算出される（ステップ15A）。この後、類似性判定装置7Aによって、類似度が類似度 α よりも低いオブジェクトが検索対象から除かれ（ステップ17A）、残ったオブジェクトについて類似度の高い値に順序付けが行われ（ステップ18A）、検索結果出力装置8によって類似度の高い順にオブジェクトが出力される（ステップ19）。

【0033】図3は本発明の第2の実施形態の類似オブジェクト検索装置の構成図、図4はその動作を示すフローチャートである。

【0034】本実施形態では類似度算出装置6Bと類似性判定装置7Bがそれぞれ第1の実施形態の類似度算出装置6Aと類似性判定装置7Aと異なっている。

【0035】すなわち、NOT指定9がなければ第1の実施形態と同じく、類似度算出装置6Bは類似度が類似度 α より低いオブジェクトを検索対象から除き（ステップ17A）、類似性判定装置7Bは残ったオブジェクトを類似度の高い順に順序付けするが（ステップ18A）、NOT指定9があれば（ステップ20）、類似度算出装置6Bは、類似度が類似度 α 以上のオブジェクトを検索対象から除き（ステップ17B）、類似性判定装置7Bは残ったオブジェクトを類似度の低い順に順序付けする（ステップ18B）。

【0036】図5は本発明の第4の実施形態の類似オブジェクト検索装置の構成図、図6はその動作を示すフローチャートである。

【0037】本実施形態では、基準オブジェクト生成装置3B、特徴抽出装置4B、類似度算出装置6C、類似性判定装置7Cがそれぞれ第1の実施形態の基準オブジェクト生成装置3A、特徴抽出装置4A、類似度算出装置6A、類似性判定装置7Aと異なっている。

【0038】すなわち、基準オブジェクト生成装置3Bは検索キーオブジェクトから複数の基準オブジェクトを生成し、特徴抽出装置4Bはこれら複数の基準オブジェクトの特徴量を抽出し（ステップ14B）、類似度算出装置6Cは検索キーオブジェクトと複数の基準オブジェクトの類似度 α_i （ $i=1\sim n$ ）を算出し（ステップ15B）、類似度の α_i （ $i=1\sim n$ ）の中で最も高い類似度 α_{max} を基準類似度として抽出し（ステップ16A）、類似性判定装置7Cは類似度が基準類似度 α_{ma}

x より低いオブジェクトを検索対象から除き（ステップ17C）、残りのオブジェクトを類似度の高い順序付けする（ステップ18A）。

【0039】図7は本発明の第4の実施形態の類似オブジェクト検索装置の構成図、図8はその動作を示すフローチャートである。

【0040】本実施形態では、類似度算出装置6D、類似性判定装置7Dがそれぞれ第3の実施形態の類似度算出装置6C、類似性判定装置7Cと異なっている。

【0041】すなわち、類似度算出装置6Dは、類似度 α_i （ $i=1\sim n$ ）の中で最も高い類似度 α_{max} と最も低い類似度 α_{min} をそれぞれ第1の基準類似度、第2の基準類似度として抽出し（ステップ16B）、NOT指定9がなければ、第3の実施形態と同じ、類似度算出装置6Dは、類似度が第1の基準類似度 α_{max} より低いオブジェクトを検索対象から除き（ステップ17C）、残りのオブジェクトを類似度の高い順に順序付けするが（ステップ18A）、NOT指定9があると（ステップ20）、類似度算出装置6Dは類似度が第2の基準類似度 α_{min} より高いオブジェクトを検索対象から除き（ステップ17D）、類似性判定装置7Dは残りのオブジェクトを類似度の低い順に順序付けする（ステップ18B）。

【0042】第1、第3の類似オブジェクト検索装置では類似しないオブジェクトの出力を抑えることが可能になり、利用者は類似するオブジェクトと類似しないオブジェクトを判定する必要がなくなる。

【0043】さらに、従来の検索では、格納済みのオブジェクトの類似度がいずれも低く、主観的に類似と思われない場合でも、類似検索装置は、類似度の順に全てのオブジェクトの出力を行っていたが、本発明を利用すれば、基準オブジェクトより類似度の高いオブジェクトが存在しない場合に「類似のオブジェクトがない」と判定することも可能になる。

【0044】また、第2、第4の類似オブジェクト検索装置では、利用者が検索キーオブジェクトの入力とともに、否定の検索を指示した場合には、類似性判定装置7B、7Dによって、その類似度が基準類似度（第2の基準類似度） α より高い場合には、類似であると判定し、検索対象から除くことによって、利用者の否定型検索を実現できる。

【0045】なお、基準オブジェクトは予め与えられたものでもよいし、特徴量格納・管理装置5に既に格納されているオブジェクトを加工または統計処理して生成してもよい。また、基準オブジェクトを検索キーオブジェクトから生成する場合を除いては、その特徴量を求めるタイミングは検索キーオブジェクト投入タイミングの前であっても構わない。

【0046】オブジェクトが画像の場合、予め準備する基準画像（基準オブジェクト）としては、何も描かれて

いない画像（白紙、黒紙、緑紙等）が容易である。検索キー画像から生成する基準画像としては、画像の画素値を平均化した画像や、画像の低周波成分のみから構成した画像や、画像の周辺の画素値から補間内挿した画像等がある。

【0047】既に格納されている画像から生成する基準画像としては、各画像を検索キー画像と見なして前記の手法を用いることや、全画像の平均画素値からなる画像等の統計量を利用した方法等がある。

【0048】音の場合、予め準備する基準音（基準オブジェクト）としては、無音や、外界の平均ノイズ音、検索キー音から生成する基準音としては、平均化した音波や、低周波成分からのみ構成した音波等がある。

【0049】テキストの場合、既に格納されているテキストから生成する基準テキスト（基準オブジェクト）としては、各テキストの平均キーワード出現数等の統計量を利用した方法等がある。

【0050】映像の場合、各シーンに対し、先頭画像や、平均画像、代表画像を作成し、画像の場合と同様の手法を利用したり、動画特有の動きベクトルを利用する方法等がある。

【0051】画像の場合には、その特徴量は、色相、輝度、彩度、エッジ、テクスチャ、構造等さまざまな基準が存在する。

【0052】ここでは、説明のため特徴量をRGBそれぞれの出現画素数とすると、特徴抽出装置4A、4Bは3つの値から成るベクトル値を生成することになる。ここでは、既に以下の5つの画像に対する特徴ベクトル値、つまり特徴量が格納されていたとする。

【0053】(0, 3, 4), (10, 4, 8), (6, 7, 8), (0, 2, 4), (8, 9, 0) さらに、キー画像（キーオブジェクト）として(6, 10, 9)が与えられたとする。類似度の基準をユークリッド距離としたとすると、キー画像と各格納画像の類似度は以下になる。なお、以下の例で、 \sqrt{x} はxの平方根、 $|x|$ はxの絶対値、 x^2 はxの2乗を表す。

【0054】(0, 3, 4)と(6, 10, 9)の類似度： $\sqrt{|0-6|^2+|3-10|^2+|4-9|^2}=\sqrt{36+49+25}=10.5$

(10, 5, 8)と(6, 10, 9)の類似度： $\sqrt{|10-6|^2+|5-10|^2+|8-9|^2}=\sqrt{16+25+1}=6.4$

(6, 7, 8)と(6, 10, 9)の類似度： $\sqrt{|6-6|^2+|7-10|^2+|8-9|^2}=\sqrt{0+9+1}=3.2$

(0, 2, 4)と(6, 10, 9)の類似度： $\sqrt{|0-6|^2+|2-10|^2+|4-9|^2}=\sqrt{36+64+25}=11.2$

(8, 9, 0)と(6, 10, 9)の類似度： $\sqrt{|8-6|^2+|9-10|^2+|0-9|^2}=\sqrt{4+1+81}=9.3$

よって、従来の方法では、(6, 7, 8), (10, 5, 8), (8, 9, 0), (0, 3, 4), (0, 2, 4)の順にキー画像に類似するということだけがかかる。

【0055】そこに本発明を適用すると、基準オブジェクト生成装置3A、3Bで生成される基準画像を、RGB各色に対する格納済み画像の平均値とすると、

$(0+10+6+0+8)/5=4.8$

$(3+5+7+2+9)/5=5.2$

$(4+8+8+4+0)/5=4.8$

(4.8, 5.2, 4.8)と(6, 10, 9)の類似度： $\sqrt{|4.8-6|^2+|5.2-10|^2+|4.8-9|^2}=\sqrt{1.44+23.04+17.64}=6.49$

となり、類似性判定装置7A、7Bでは類似画像は6.49より小さいものであると判定でき、(6, 7, 8), (10, 5, 8)の2画像だけを類似画像とし、検索結果出力装置8により出力する。

【0056】また、類似性判定装置7B、7Dでは6.49より大きいものは類似でない画像と判定することもできるので、検索条件として否定条件NOT9が指定されていた場合には、類似性判定装置7B、7Dでは(8, 9, 0), (0, 3, 4), (0, 2, 4)を類似でない画像とし、検索結果出力装置8により出力する。

【0057】同様に、基準オブジェクト生成装置3A、3Bにおいて、値域の最大値(10)や中間値(5)を基準画像とした場合には以下のような類似度になり、これらによって類似性の判定を行うことができる。

【0058】(10, 10, 10)と(6, 10, 9)の類似度： $\sqrt{|10-6|^2+|10-10|^2+|10-9|^2}=\sqrt{16+0+1}=4.1$

(5, 5, 5)と(6, 10, 9)の類似度： $\sqrt{|5-6|^2+|5-10|^2+|5-9|^2}=\sqrt{1+25+16}=6.5$

さらに、上記3つの基準画像を同時に利用した場合には、類似度が $\min(6.49, 4.1, 6.5)=4.1$ より小さいものを類似であると判定するため、(6, 7, 8)だけが類似である画像となる。逆に、類似度が $\max(6.49, 4.1, 6.5)=6.5$ より大きいものは類似でないと判定するため、(8, 9, 0), (0, 3, 4), (0, 2, 4)が類似でない画像となる。

【0059】

【発明の効果】以上説明したように、本発明は、類似オブジェクト検索において、基準オブジェクトを用いた判

定処理を用いることにより、「類似する」と「類似しない」の分離が可能になったことで以下のような効果が得られる。

【0060】・検索結果からノイズ（類似しないもの）を自動的に除くことができ、検索精度が向上し、利用者の判断の負担が減少し、同時に検索速度も向上する。

【0061】・否定形（類似しない）の問い合わせの実現が可能になり、「検索キー画像Xに似ており、かつ、検索キー画像Yには似ていない」のような肯定条件と否定条件を組み合わせた問い合わせも可能になる。

【0062】・これは基準オブジェクトとの類似度より小さい類似度を持つ格納済み画像がない場合には「該当結果なし」と判定することが可能になる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態の類似オブジェクト検索装置の構成図である。

【図2】第1の実施形態の動作を示すフローチャートである。

【図3】本発明の第2の実施形態の類似オブジェクト検索装置の構成図である。

【図4】第2の実施形態の動作を示すフローチャートである。

【図5】本発明の第3の実施形態の類似オブジェクト検

索装置の構成図である。

【図6】第3の実施形態の動作を示すフローチャートである。

【図7】本発明の第4の実施形態の類似オブジェクト検索装置の構成図である。

【図8】第4の実施形態の動作を示すフローチャートである。

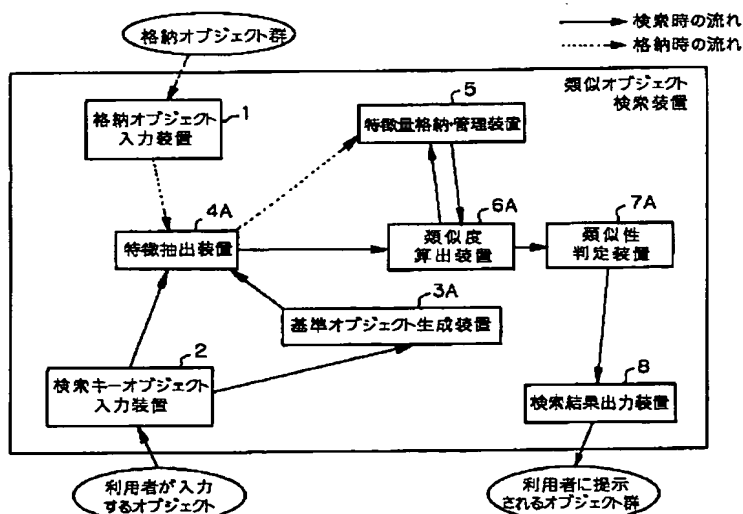
【図9】類似オブジェクト検索装置の従来例の構成図である。

【図10】図9の従来例の動作を示すフローチャートである。

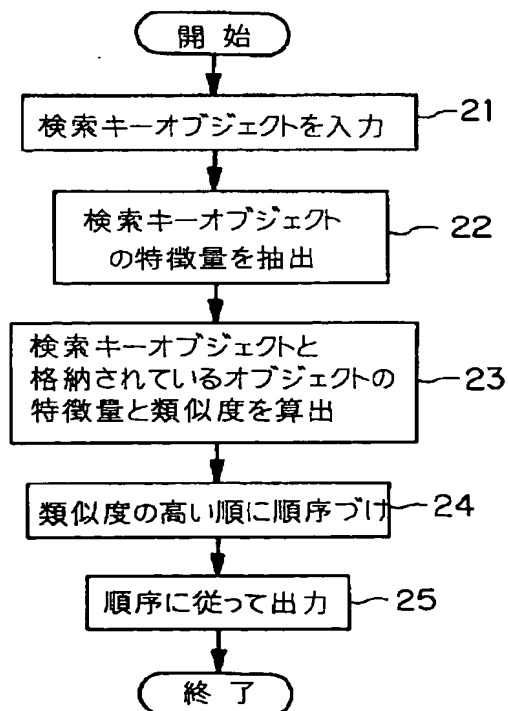
【符号の説明】

- 1 検索オブジェクト入力装置
- 2 検索キーオブジェクト入力装置
- 3 A, 3 B 基準オブジェクト生成装置
- 4 A, 4 B, 4 C 特徴量抽出装置
- 5 特徴量格納・管理装置
- 6 A, 6 B, 6 C, 6 D, 6 E 類似度算出装置
- 7 A, 7 B, 7 C, 7 D, 7 E 類似性判定装置
- 8 検索結果出力装置
- 9 否定入力
- 11～20 ステップ

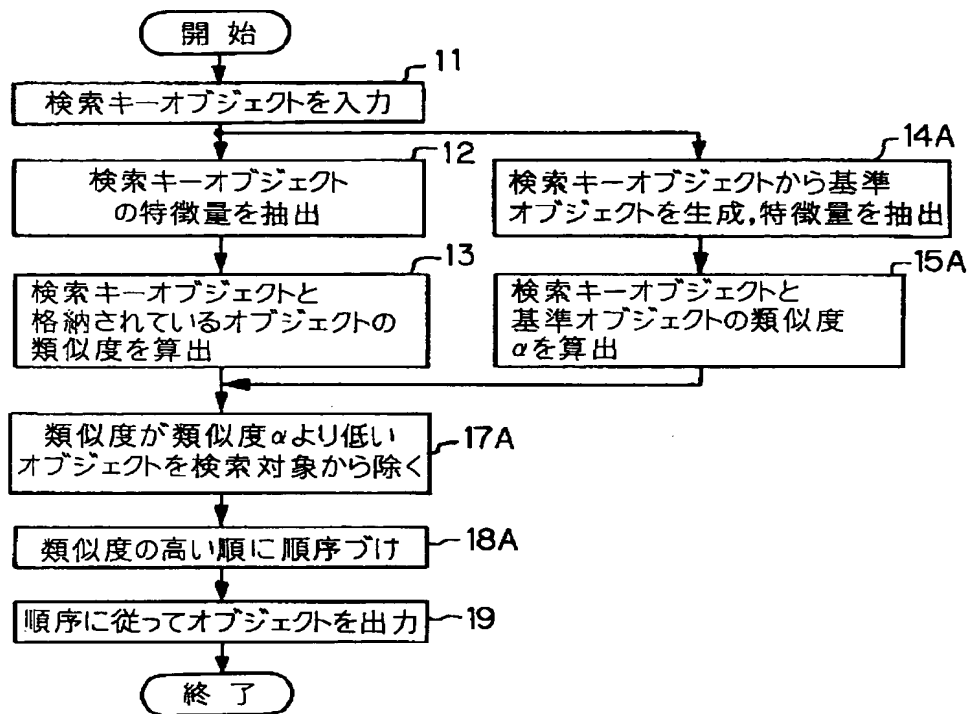
【図1】



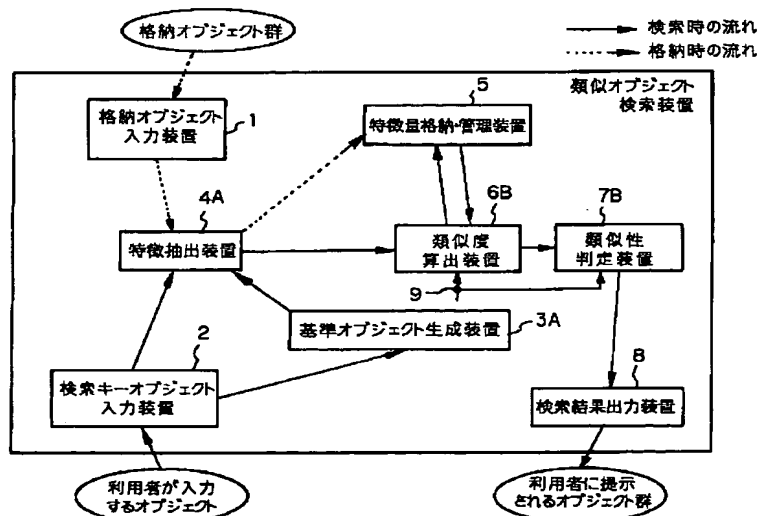
【図10】



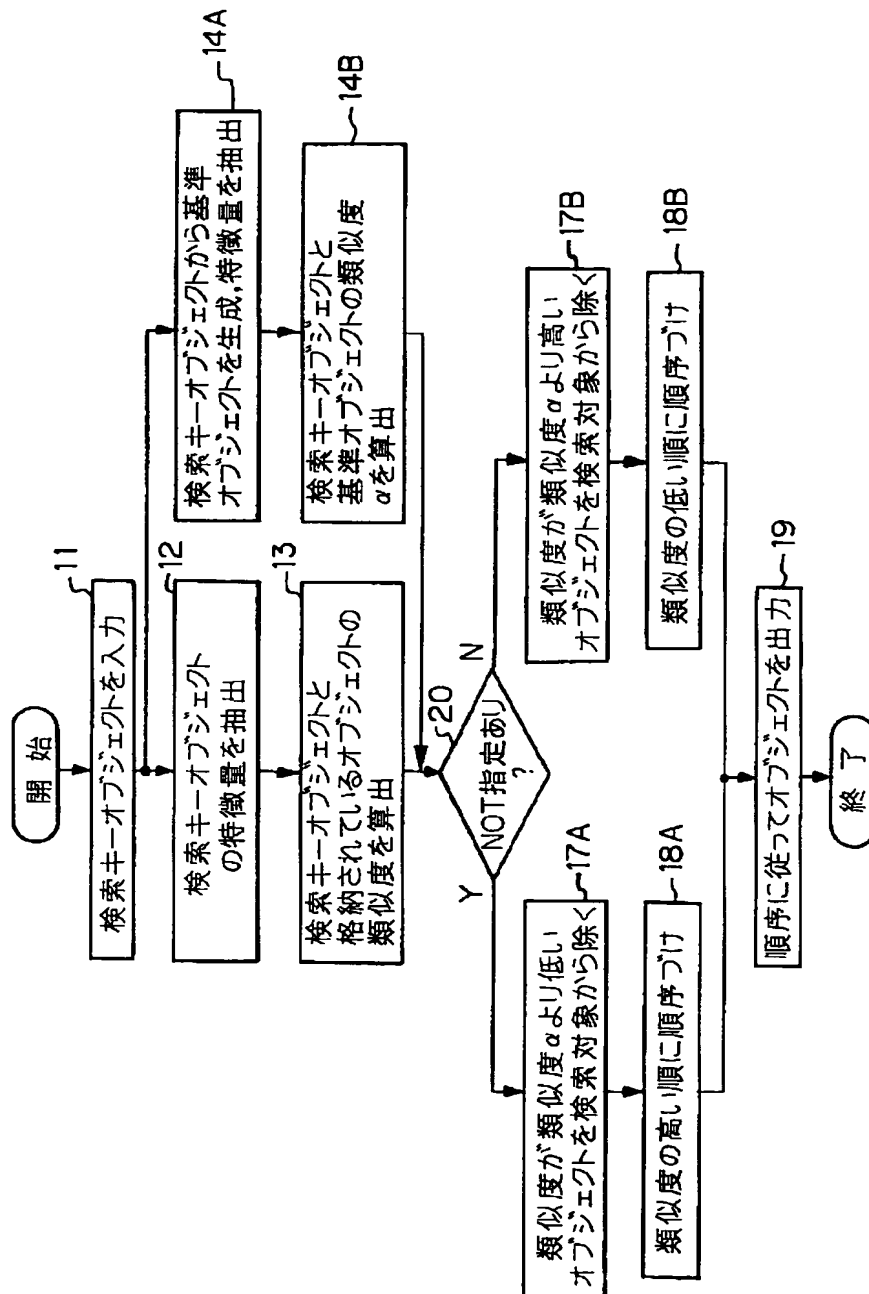
【図2】



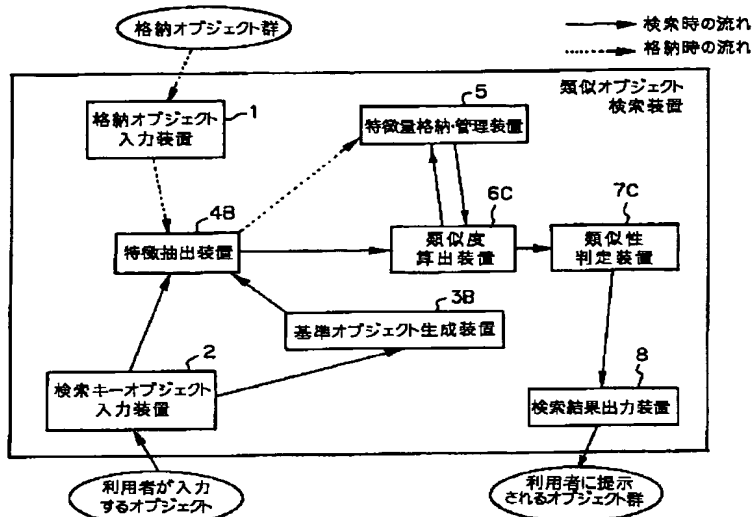
【図3】



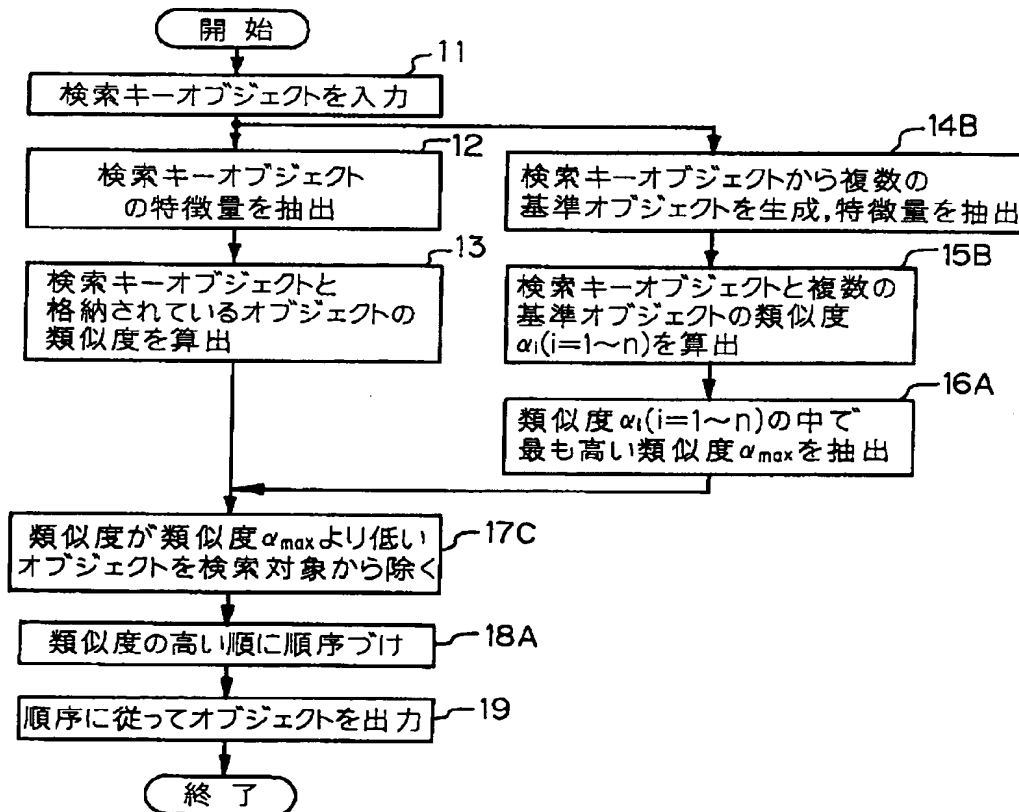
【図4】



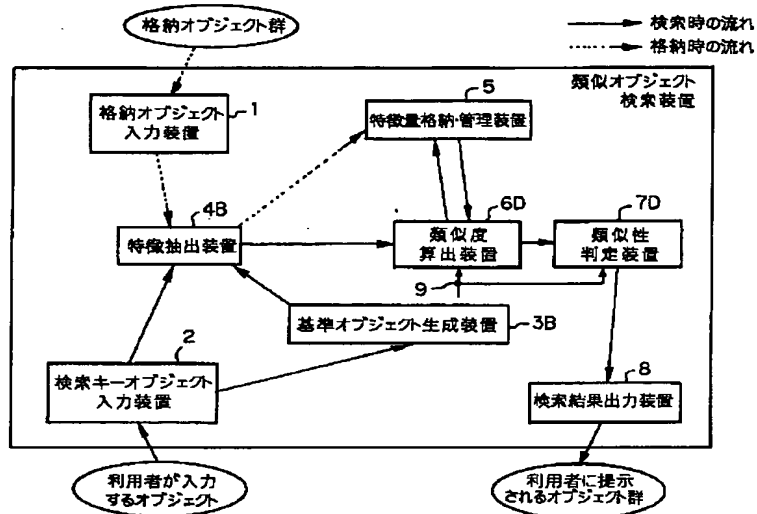
【図5】



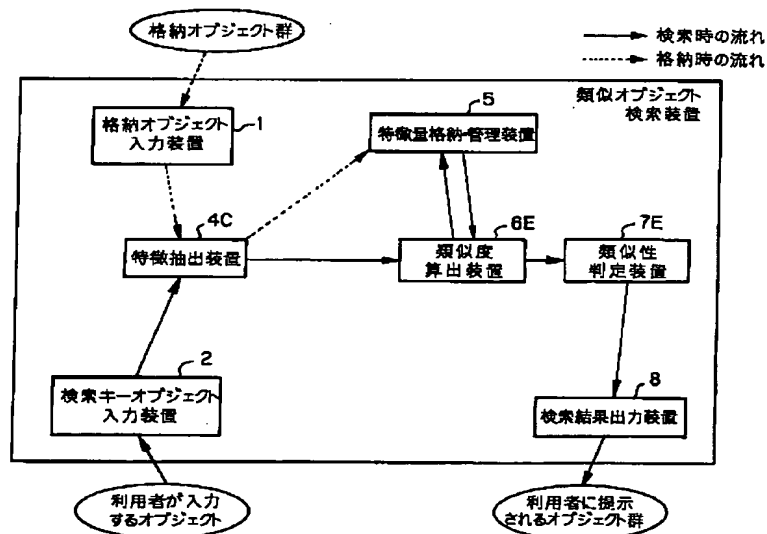
【図6】



【図7】



【図9】



【図8】

